

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
3
3
3
3
3
3
3
3
3
3

5
6
7

8

9
10
11
12
13
14

15
16
17
18
19
20
21
22
23
24

25
26

27
28
29
30
31
32
33

3.
3.
3
3
3

1 DAMP sending client; and setting a destination IP address field to the non-zero IP address
2 of one of the selected multiple remote destinations.

3 In yet another respect, what is described is a computer-readable medium on which
4 is embedded a program. The embedded program includes instructions for executing the
5 above method.

6 Those skilled in the art will appreciate these and other advantages and benefits of
7 various embodiments of the invention upon reading the following detailed description of
8 an embodiment with reference to the below-listed drawings.

9 10 BRIEF DESCRIPTION OF DRAWINGS

11 Figure 1 is a block diagram of a system for delivering multicast data packets over
12 a network according to the prior art;

13 Figure 2 is a block diagram of one embodiment of a system for transmitting data
14 packets over a network to selected multiple remote destinations according to the present
15 invention;

16 Figure 3 is a block diagram of another embodiment of a system for transmitting
17 data packets over a network to selected multiple remote destinations according to the
18 present invention;

19 Figure 4 is a flowchart of one embodiment of a method for transmitting data
20 packets over a network to selected multiple remote destinations;

21 Figure 5 is a diagram of one embodiment of the structure of a directly addressed
22 multicast protocol data packet according to the present invention;

23 Figure 6 is a diagram of one embodiment of the structure of an IP options field
24 within a directly addressed multicast protocol data packet according to the present
25 invention;

26 Figure 7 is a diagram illustrating a path of one embodiment of a directly
27 addressed multicast protocol data packet through the Internet to multiple destinations; and

28 Figure 8 is a diagram illustrating another path of one embodiment of a directly
29 addressed multicast protocol data packet through the Internet to multiple destinations.

30 31 DETAILED DESCRIPTION

32 The present invention relates to a new IP multicast protocol hereinafter referred to
33 as Directly Addressed Multicast Protocol (DAMP). In a DAMP system, a sending client
34 specifies a list of remote destinations directly for each data packet being sent. The term
35 DAMP is used as a label only, and other terms can define the same or similar protocols.

36 Using DAMP, a network of switches and routers will continue to resend the data
37 packets over only those network segments that contain a route to at least one of the
38 specified remote destination addresses. This method thus does not generate unnecessary

1 traffic on any network segment. This form of multicast also greatly simplifies the
2 multicast process by removing the need for multicast group management protocols to
3 manage the destination lists within the network infrastructure. The list of remote
4 destinations are embedded in each data packet, rather than being persistent within the
5 network infrastructure.

6 Figure 1 is a block diagram of a system 100 for delivering multicast data packets
7 over a network according to the prior art. In the existing art covering IP multicast
8 protocols, the system 100 includes at least one IP multicast sending client 110 connected
9 to and communicating with a network infrastructure 115. The IP multicast sending client
10 110 sends IP multicast data packets out to the network infrastructure 115. Within the
11 network infrastructure 115, at least one IP multicast subscription manager 120 is further
12 connected to and communicating with a number of network devices 140. The IP
13 multicast subscription manager 120 receives requests from the network devices 140 to
14 subscribe to multicast data transmissions from the IP multicast sending client 110. When
15 the IP multicast sending client 110 sends IP multicast data packets out to the network
16 infrastructure 115, the IP multicast subscription manager 120 determines which of the
17 network devices 140 have subscribed to those particular data packets. Alongside the IP
18 multicast subscription manager 120 within the network infrastructure 115 are a number of
19 network switches and routers 130. Upon a determination of which network devices 140
20 are subscribed to the data packets being transmitted from the IP multicast sending client,
21 the network switches and routers 130 forward copies of the data packets to each of the
22 network devices 140 subscribed to the data packets, and await receipt acknowledgments
23 from the network devices 140. The network infrastructure 115 may resend data packets
24 until all data packets are confirmed received by all network devices 140. As a result of
25 this multicast scheme, the data packets may travel over any given network segment many
26 times, multiplying the load on the network bandwidth.

27 Figure 2 is a block diagram of one embodiment of a system 200 according to the
28 present invention for transmitting data packets over a network to selected multiple remote
29 destinations. The system 200 of the present invention includes a DAMP sending client
30 210 which is connected to one or more network elements 220. The DAMP sending client
31 210 transmits DAMP data packets to the network elements 220. A network element 220
32 may be a network switch or router and is a part of the overall network infrastructure 115,
33 as described for Figure 1. However, unlike existing multicast systems 100, the present
34 invention does not require an IP multicast subscription manager 120 to manage requests
35 for subscriptions to multicast data transmissions and to manage lists of requesting
36 network devices 140. The system 200 of an embodiment of the present invention
37 includes a number of remote network devices 240. In the present invention, a list of
38 selected remote network devices 240 that are to receive the DAMP multicast transmission

1 is embedded in the DAMP data packets themselves. The network elements 220 route the
2 DAMP data packets on to the selected remote network devices 240 determined in the list
3 embedded in the DAMP data packets. In the system 200, it is not necessary for the
4 DAMP data packets to travel over any given network segment more than once, thus
5 reducing the load on the bandwidth of the network.

6 Figure 3 is a block diagram of another embodiment of a system 300 for
7 transmitting data packets over a network to selected multiple remote destinations
8 according to the present invention. The system 300 illustrates the use of the present
9 invention in a network storage application. The system 300 includes a DAMP sending
10 client 310 connected to one or more network elements 320. The DAMP sending client
11 310 transmits DAMP data packets to the network elements 320. As with the network
12 elements 220 of the system 200, the network elements 320 may include network switches,
13 routers, or other network infrastructure elements. The network elements 320 are
14 connected to a number of remote network storage devices 340. As with the embodiment
15 described for Figure 2, in the system 300, a list of selected remote network storage
16 devices 340 that are to receive the DAMP multicast transmission is embedded in the
17 DAMP data packets. The network elements 320 route the DAMP data packets on to the
18 selected remote network storage devices 340 determined in the list embedded in the
19 DAMP data packets.

20 Figure 4 is a flowchart of one embodiment of a method 400 for transmitting data
21 packets over a network to selected multiple remote destinations. The method 400
22 includes the steps of: embedding a list of remote destination addresses into data packets
23 (step 420); enabling network elements to access the list of remote destination addresses
24 (step 430); and instructing network elements to transmit copies of the data packets to each
25 address on the list of remote destination addresses (step 440).

26 In one embodiment of the method 400, the embedding step 420 includes the
27 additional steps of: setting up an IP Options field within the IP header section of a
28 DAMP data packet; setting a Code byte within the IP Options field to a specific value to
29 indicate that the data packet is a DAMP data packet; setting a Length byte to a
30 determinable value to indicate the length in 32-bit words of the IP Options field;
31 embedding in successive 32-bit words the values of a determinable number of IP
32 addresses for the multiple remote destinations for the DAMP data packet; and setting the
33 source IP address in the header section of the DAMP data packet to the IP address of the
34 DAMP sending client 210; and setting the destination IP address in the IP header section
35 of the DAMP data packet to the IP address of one of the multiple remote destinations
36 embedded in the IP Options field. In this embodiment of the method 400, the copying
37 step 440 further includes the steps of: the network element 220 receiving a copy of the
38 DAMP data packet from the DAMP sending client 210 or from another network element

1 220; the network element 220 zeroing those IP addresses embedded in the IP Options
2 field that are not directly accessible below the network element 220; the network element
3 220 setting the destination IP address in the IP header section of the DAMP data packet to
4 the IP address of one of the non-zeroed multiple remote destinations embedded in the IP
5 Options field; and the network element 220 routing a copy of the modified DAMP data
6 packet to each additional network element 220 or network device 240 for which there is a
7 corresponding non-zeroed IP address listed in the embedded list of multiple remote
8 destination IP addresses.

9 Figure 5 is a diagram of the structure 500 of one embodiment of a DAMP data
10 packet 510 according to one embodiment of the present invention. The data packet
11 structure 500 includes a data packet header section 515 and a data packet data section
12 525.

13 The data packet header section 515 contains multiple fields, one of which is a
14 variable length field containing certain IP options information, an IP options field 520.
15 Figure 6 is a diagram of one embodiment of the structure of an IP options field 520 within
16 a directly addressed multicast protocol data packet 510 according to the present invention.
17 The variable length IP options field 520 comprises a sequence of items, each of which
18 start with a Code byte 610 comprised of the following bits, where bit 0 is considered the
19 Most Significant Bit ("MSB"):

20 Bit 0 - Copy bit - which defines whether the IP options field 520 should be copied
21 into each network fragment if the data packet 510 is fragmented, or split, across
22 multiple network frames. A value of 0 indicates that the IP options field 520
23 should be copied into only the first frame, whereas a value of 1 indicates that the
24 IP options field 520 should be copied into every frame;

25 Bits 1-2 - Option class - which defines a set of IP class values;

26 Bits 3-7 - Option number - these bits identify the specific IP option for this data
27 packet, where each of the available IP options is associated with a unique Option
28 number.

29 An IP data packet 510 may be encoded with DAMP using an IP option Code byte
30 610 which in one embodiment may be set to a value of 138. Other values for the IP
31 option Code byte 610 may be equally possible. A Code byte 610 value of 138
32 corresponds to setting the Copy bit to 1, the Option class to 0, and the option number to
33 10, where 10 is one of the currently unused values for IP option numbers. The Option
34 number may alternatively be any currently unused IP option value.

35 Following the Code byte 610 in the encoding of the DAMP data packet IP options
36 field 520, the structure of a DAMP data packet header section 515 may follow a pattern
37 similar to that of the existing Strict Route IP Option, which is currently assigned to IP
38 Option number 137. See Comer, Douglas E., Internetworking with TCP/IP - Principles,

1 Protocols, and Architecture, 4th Ed., Vol. 1, Prentiss Hall, February 2000, pp.97-114.

2 The IP options field 520 is broken down into multiple 32-bit words, as shown in
 3 Figure 6. The first word of the IP options field 520 contains the 8-bit (one byte) Code
 4 byte 610, set to a value of 138 for DAMP data packets, and the Length byte 615, which
 5 contains a value specifying the number of bytes in the IP options field 520 (including the
 6 code and length bytes 610 and 615). This word is followed by multiple IP address fields
 7 620, each containing four-byte IP addresses for each of the intended recipients of the
 8 DAMP data packet 510, the remote network devices 240. There may be as many IP
 9 address fields 620 as allowed for by the value specified in the Length byte 615. For
 10 example, if the Length byte 615 indicates that the IP options field 520 includes 32 bytes,
 11 then the IP options field 520 contains seven 32-bit (four-byte) IP address fields 620, in
 12 addition to the first word containing the Code byte 610 and the Length byte 615. The
 13 data packet header section 515 may also include two other fields of interest: a source IP
 14 address field 522, specifying the IP address of the DAMP sending client 210, and a
 15 destination IP address field 524, specifying the IP address of one of the remote network
 16 devices 240. In an alternative embodiment, the IP address fields 620, the source IP
 17 address field 522, and the destination IP address field 524 may each comprise multi-byte
 18 IP addresses of a type other than four-byte addresses.

19 The data packet data section 525 is encoded following a standard User Datagram
 20 Protocol ("UDP") IP data packet encoding scheme. As a connectionless protocol which,
 21 like TCP, is layered on top of IP, UDP neither guarantees delivery nor does it require a
 22 connection. As a result, it is lightweight and efficient, but all error processing and
 23 retransmission must be taken care of by the application program. See Postel, Jon, User
 24 Datagram Protocol, RFC 768, Network Information Center, SRI International, Menlo
 25 Park, Calif., August 1980.

26
 27 Figure 7 shows an exemplary path of one embodiment of a DAMP encoded data
 28 packet 711 through the Internet to multiple destination remote network devices 240.
 29 When a DAMP sending client 210 sends a DAMP data packet 711, it sets the source IP
 30 address 522 contained in the data packet header section 525 to the IP address of the
 31 DAMP sending client 210 itself, and the destination IP address 524 to any of the IP
 32 addresses of the remote network devices 240 intended to receive the data packet. The
 33 DAMP sending client 210 further encodes the data packet 711 with an IP Options field
 34 520 containing the value for the DAMP IP Option, as described above, and which
 35 contains a list of each of the desired destination remote network device 240 IP addresses.
 36 The data packet data section 525, or UDP-encoded data portion of the data packet, thus
 37 remains unchanged from that of a standard UDP data packet which may be sent to any
 38 single recipient under existing IP transmission protocols.

1 When a router (or network switch) 715 receives a DAMP data packet 711, which
 2 a router or switch may recognize from the DAMP IP Option encoded into the data packet
 3 711, it will ignore the destination IP address 524, and instead examine the list of IP
 4 addresses contained in the DAMP IP Options field 520. The router 715 then sends a copy
 5 to each network interface 720 and 730 that contains at least one recipient as specified in
 6 the address list in the IP Options field 520. However, before sending the data packets
 7 721 and 731 on to a network interface, such as the network interfaces 720 and 730 shown,
 8 each IP address in the IP Options field 520 list that is not intended to eventually receive
 9 that copy of the data packet, i.e., is not found on a branch of that specific network
 10 interface, is zeroed out. This prevents the creation of an infinite number of data packets
 11 sent by two interconnected routers or switches addressing hosts existing across more than
 12 one network interface. Furthermore, the switch or router 715 will set the destination IP
 13 address 524 in the IP header section 515 to one of the non-zero entries remaining in the
 14 recipient list (and the corresponding frame header destination hardware address). When a
 15 DAMP data packet 721 or 731 is received with zeroed addresses in it, the zeroed
 16 addresses are ignored. Entries are zeroed out rather than removing them in order that the
 17 router or switch does not have to reformat the data packet.

18 Figure 8 is a diagram showing how a network of DAMP enabled routers and
 19 switches would pass a DAMP data packet through several branches of the network.
 20 Exemplary values for the relevant fields embedded in each successive copy of the DAMP
 21 data packet are shown as it travels across each network segment. This figure shows how
 22 in one embodiment of the invention, a network element 220, such as a router or IP switch
 23 (815, 825, 835, 837), upon receiving a copy of the DAMP data packet 510 (811, 821,
 24 831, 832) from the DAMP sending client 210 or from another network element (815, 825,
 25 835, 837), then processes a copy of the received DAMP data packet (811, 821, 831, 832)
 26 by zeroing those IP addresses embedded in the IP Options field 520 that are not directly
 27 accessible below the network element (815, 825, 835, 837); setting the destination IP
 28 address 524 in the IP header section 525 of the copy of the DAMP data packet (821, 822,
 29 831, 832, 841, 842, 843) to the IP address of one of the non-zeroed remote destination
 30 network devices (851, 852, 853, 854) embedded in the IP Options field 520; and routing
 31 the modified copy of the DAMP data packet (821, 822, 831, 832, 841, 842, 843) to each
 32 additional network element (825, 835, 837) or network device (851, 852, 853, 854) for
 33 which there is a corresponding non-zeroed IP address listed in the embedded list of
 34 multiple remote destination IP addresses.

35 When a remote network device 240 receives a DAMP data packet, it can process
 36 it like any other UDP data packet. In fact, no DAMP software is necessary on the remote
 37 network devices 240 beyond the normal UDP software necessary to receive the data
 38 packet. In an alternate embodiment, the remote network device's 240 network element

1 220 could be programmed or otherwise adapted to be DAMP aware such that the network
2 element 220 could examine the list of addresses contained in the DAMP IP Options field
3 520 to determine if the DAMP data packet 510 was intended for that network element
4 220. This would remove the need for switches and routers to set the destination IP
5 address 524 to a recipient IP address (which may be impractical for network elements 220
6 to do without duplicating the data packet 510 and sending a copy of the data packet 510
7 for each recipient address to every network device 240 attached to it).

8 An alternative embodiment of DAMP data packet encoding may store destination
9 port numbers within the recipient list along with the IP addresses. Routers and switches
10 could then set the port numbers within the UDP header in the IP data section of the
11 DAMP data packet. This alternative embodiment may eliminate one potential limitation
12 of the DAMP encoding scheme whereby it is otherwise assumed that the destination port
13 number is identical for each recipient host.

14 The method 400, shown in Figure 4, is better understood in conjunction with the
15 diagram of the structure 500 of a DAMP data packet 510, as shown in Figure 5, and the
16 diagram of one embodiment of the structure of an IP options field 520 within a directly
17 addressed multicast protocol data packet 510, as shown in Figure 6.

18 The method 400 operates by embedding (step 420) multiple IP address fields 620
19 into the header section 515 of a DAMP data packet 510. Additionally, the method 400
20 provides a scheme (step 430) by which network elements 220, such as network routers or
21 switches, may access the list of multiple IP address fields 620. One embodiment of the
22 present invention places a uniquely formatted IP options field 520 in the IP header section
23 515 of the DAMP data packet 510. The IP options field 520, by way of a specific value
24 assigned to a Code byte 610 within the IP Options field 520, signifies that the data packet
25 510 is a DAMP data packet and thereby indicates that a list of multiple IP address fields
26 620 is embedded in the IP header section 515 of the DAMP data packet 510. The IP
27 options field 520 may then be read and translated (step 430) by network elements 220 and
28 the list of multiple IP address fields 620 embedded in the IP header section 515 of the
29 DAMP data packet 510 may then be accessed and read. The network elements 220 then
30 transmit (step 440) copies of the DAMP data packet 510 to each of the addresses in the
31 list of multiple IP address fields 620.

32 The present invention improves on existing multicast protocols in one respect by
33 not consuming bandwidth on a network segment to send the same data twice. This results
34 in lower network utilization and improved network performance. DAMP further
35 improves on existing multicast protocols by not requiring the overhead of setting up and
36 deleting multicast groups in the network switch, independent of the data packets, and by
37 not requiring a subscription service for the transmission of multicast data. This also
38 results in lower network utilization and improved performance.

1 For one embodiment of the present invention, a comparison may be made, outside
2 the unique DAMP addressing model and multicast behavior, between the network effects
3 of DAMP and the well known Unreliable Datagram Protocol ("UDP"), in contrast to the
4 Transmission Control Protocol ("TCP"). This analogy may be drawn where the
5 embodiment of the present invention does not require confirmation of receipt of the
6 DAMP data packets by the remote network devices to be sent back to the DAMP sending
7 client before continuing with transmission of additional DAMP data packets, as in UDP.
8 This embodiment substantially reduces the load on the network.

9 In another embodiment of the present invention, confirmation of receipt of the
10 DAMP data packets by the remote network devices is required to be sent back to the
11 DAMP sending client, similar to the manner in which TCP operates. This embodiment
12 requires bidirectional transmission of data, resulting in a correspondingly lesser decrease
13 in network traffic than the embodiment described above.

14 One potential use identified for DAMP is in networked storage systems. DAMP
15 allows data clients to send data directly to storage elements, and enables the data clients
16 to arrange the data in striped and mirrored configurations across numerous remote storage
17 elements, each with their own unique IP addresses, without the need for multicast
18 subscription management elements within the network infrastructure. DAMP allows the
19 client to efficiently and dynamically, with no unnecessary overhead, deliver data to the
20 appropriate remote storage elements.

21 The steps of the method 400 can be implemented with hardware or by execution
22 of programs, modules or scripts. The programs, modules or scripts can be stored or
23 embodied on one or more computer readable mediums in a variety of formats, such as
24 source code, object code or executable code, for example. The computer readable
25 mediums may include, for example, both storage devices and signals. Exemplary
26 computer readable storage devices include conventional computer system RAM (random
27 access memory), ROM (read only memory), EPROM (erasable, programmable ROM),
28 EEPROM (electrically erasable, programmable ROM), and magnetic or optical disks or
29 tapes. Exemplary computer readable signals, whether modulated using a carrier or not,
30 are signals that a computer system hosting or running the described methods can be
31 configured to access, including signals downloaded through the Internet or other
32 networks.

33 The terms and descriptions used herein are set forth by way of illustration only
34 and are not meant as limitations. Those skilled in the art will recognize that many
35 variations are possible within the spirit and scope of the invention as defined in the
36 following claims, and their equivalents, in which all terms are to be understood in their
37 broadest possible sense unless otherwise indicated.

38